A COMPUTER-BASED SYSTEM FOR INDEXING CURRICULUM CONTENT

Judith L. Piggins*, Harvey Goldman MD**, Diane Oliver MD*, Dyan Ryan Blewett*, Laurie J. Hassan*, Alane Shanks**, Susan Zapatka**, and G. Octo Barnett MD*

*Laboratory of Computer Science
Massachusetts General Hospital
Boston, MA 02114

**Office of Educational Development
Harvard Medical School
Boston, MA 02115

Abstract

Harvard Medical School has developed a computer-based system for tracking information about the medical school curriculum. Based on five years of experience with monitoring curriculum content, the system is designed for ease of use and supports searches on a number of general topic areas. These searches are then refined by review of the fields of the individual records retrieved. To facilitate the capture of data from faculty authors, the system produces a questionnaire which lists all fields and the vocabularies specified for them in a checklist format. The system is described and the lessons learned during its development and use are discussed.

Introduction

Over the past decade Harvard Medical School (HMS) has undertaken a basic restructuring of its curriculum. This experiment, called the New Pathway program [1,2], involves a movement away from rote memorization toward an increased emphasis on problem-solving and independent learning. To help accomplish this, the traditional teaching format of the first two years of medical school has been modified in a number of ways, including a decrease in the number of lectures and the introduction of problem-based tutorials. The experience gained from the new curriculum is used to refine both course content and curriculum structure on an ongoing basis.

Another key element of the New Pathway curriculum is the emphasis on information technology as an important resource for student learning and for support of faculty and staff in development of the new curriculum [3]. Thanks to a generous grant from Hewlett Packard Corporation, HMS has had access to a sizable amount of computer hardware as well as funding for staff to develop software and support users in the New Pathway program [4]. The HMS Information Technology staff is based at the Laboratory of Computer Science (LCS) at Massachusetts General Hospital. Staff from the Office of Educational Development (OED) at HMS (the headquarters for curriculum planning) collaborated with LCS staff to develop a computer-based tracking system for curriculum content.

Need for the System

A computer-based curriculum indexing system is of key importance in our environment for several reasons. First, the frequent changes in the curriculum raise the possibility that some essential topics will not be covered while others may be presented redundantly. A database of curriculum content makes it possible for faculty and curriculum planners in one subject area to easily identify the content being covered by those in other disciplines. Any deficiencies or redundancies can then be corrected. Also, the learning issues for each session are less obvious in a case-based format than in a traditional lecture-based curriculum. The syllabus for a lecture-based course generally provides an outline of the content of the course. In contrast, case titles often give little information about the topics touched on in studying the cases. A database is needed to store and retrieve the educational content information provided by authors submitting new cases. Finally, we needed a system that would allow us to retrieve data from the database efficiently and in a variety of ways, many of which might not be foreseen at the outset. All of these factors argued for a computer-based system rather than simple paper and pencil tracking.

History

The first pilot group of New Pathway students entered Harvard Medical School in September of 1985. The initial version of a curriculum tracking system, called the Case-Objectives Database, was installed in November, 1985. This system was based on the FileMan database developed by the Veterans' Administration hospitals [5]. FileMan is written in the MUMPS [tm] language

0195-4210/90/0000/0210$01.00

[6]. The system ran on an HP150 personal computer using DTM-PC MUMPS.

FileMan provides a hierarchical structure which was used to organize the curriculum information according to the subject matter blocks in which it was taught. In addition to listing the title, date and type of each session and the faculty person in charge of it, both keywords and free text could be entered to describe its educational objectives. Wherever possible, fields were restricted to a controlled vocabulary of terms in order to facilitate later searches of the database. The keywords describing educational objectives were selected from a subset of the Medical Subject Headings (MeSH) vocabulary developed by the National Library of Medicine to index the medical literature [7]. This vocabulary was chosen because it is the most generally accepted vocabulary describing the nomenclature of medicine. Other schools which have developed systems to index their curricula [8,9, 10] have generally used the MeSH vocabulary or a subset of it.

During the next three years, staff from OED and LCS worked together to try to address the problems and limitations that arose as the Case-Objectives system was put to use. The major issues encountered were as follows:

1) Searching the database required a detailed knowledge of the structure of the files and generally required the assistance of the programmer who set up the system. Although predefined templates were created to do a number of standard searches, any modifications to these searches or ad hoc requests required technical consultation.

2) The data entry sequence proved time-consuming and difficult to use for this application.

3) The data fields for each teaching session provided in the original system proved inadequate to allow the entry of all desired information, particularly for the case/tutorial sessions. Additional fields were defined to accommodate items such as presenting complaint, drugs and lab tests.

4) We found that the MeSH vocabulary was not particularly well-suited to indexing curriculum content in our environment. Choosing the correct term from MeSH requires a background in medical science which our data entry person did not have, so this task was generally assigned to a medical student. However, there was not always a student available to carry out this task, which meant that the keyword indexing often fell well behind the rest of the data entry. Also, MeSH provides a hierarchy of terms, increasing in detail as the depth of the tree increases. Different students often selected different levels of specificity for the cases they keyworded, so the encoding was not uniform. Finally, MeSH includes few terms in areas such as signs and symptoms or pathophysiology terms, and does not always include terms describing general topics such as bilirubin metabolism or genetic defects.

Current System

In February, 1990 a new curriculum-tracking system was installed at OED. Called Curriculum Index, this system represents efforts to address the issues that arose during use of the earlier version of the system. The new system runs on an HP Vectra personal computer under DTM-PC MUMPS. The design criteria used in developing Curriculum Index were as follows:

1) The system should be simple to use, both for entry of data and for retrieval. It should support routine use by physicians and other staff without the need for consultation with a programmer.

2) The number of fields should be kept small, representing broad general categories of information rather than specific groupings. For instance, we defined a single field called **Anatomical Structure** rather than separate fields for topics such as **Organs**, **Tissues**, **Structures** and **Cells**. This allows searches to be done by specifying one or more general criteria; the matching records found can then be reviewed in more detail.

3) To facilitate ease and accuracy of retrieval, as many fields as possible should have controlled vocabularies. This limits the items that can be entered into a field to a prespecified list. It avoids issues such as typographical errors, use of synonyms or use of more specific or less specific terms which can result in matching records not being correctly identified.

4) The level of detail of encoding in the controlled vocabularies should be fairly general. For example, associating a teaching session with the general term "cardiovascular system" rather than more specific terms such as "pulmonary valve" or "aorta" forces both encoding and retrieval to be done at the same high level. This ensures that searches will find all appropriate records.

5) Capture of information about each case or other session should be done via a paper form filled out by the faculty author at the time the course material is submitted. To make this practical, all controlled vocabularies should be as concise as possible to allow them to be presented in a brief checklist format that can be filled out rapidly.

Curriculum Index is based on a general-

purpose MUMPS database system called INFOMAN developed by LCS. INFOMAN provides a screen-oriented environment for data entry. A sample data entry screen from the Curriculum Index system is shown in Figure 1. The current version of the system provides a single level of data organization rather than a hierarchical format. Fields may be specified as single-entry, multiple-entry or free-text fields. Each field can be customized as to how its data may be entered. For instance, entries may be restricted to items chosen from a controlled vocabulary, or the format of the entry may be prespecified (e.g. "last name, first name"). All of these specifications can be easily modified through INFOMAN, as can the format of the fields as they are displayed on the data entry screens. This flexibility allowed us to develop the system by creating an initial version and refining it in stages.

Data may be entered for each case tutorial, lab, conference and lecture. The data items allowed for each session are listed in Figure 2. An effort was made to use controlled vocabularies wherever possible to standardize the data and enhance retrieval capabilities. The MeSH vocabulary was not used, since, as discussed above, we found it to be inappropriate for this application. Instead, we developed vocabularies of our own for each field. Although many of the terms used are MeSH terms, that is not a requirement. An effort was made to keep these vocabularies as concise as possible, both to simplify the data entry process and to make it practical to provide checklists to faculty authors to indicate the entries to be made for cases they have written.

The active participation of one of the authors (Goldman) in his role as Faculty Dean for Medical Education and Chairman of the HMS Curriculum Committee was important in providing end-user input into the creation of the system. Doctors Goldman, Oliver and Barnett collaborated in the creation of the controlled vocabularies to ensure that the lists of terms adequately indexed the entire HMS curriculum.

The largest controlled vocabulary at present is for the **Anatomical Structure** field; it includes 57 terms. Fields such as **Lab Tests** specify categories of tests rather than individual tests in the interest of brevity. For the **Diagnosis** field, where the potential number of terms is very large, we did not specify a controlled vocabulary but instead made this a free-text field. This field can be searched by specifying either an exact match to the contents of the field or a substring located anywhere within the field. The user can also request a list of the diagnoses that have been entered in existing records when choosing terms on which to search. We feel that the advantages of being able to capture the data on paper forms and being able to enter it without the constraints of a large and cumbersome controlled vocabulary outweigh the drawbacks of possible data entry errors or differences in the detail level in entries for this field.

One of the more frustrating aspects of the original Case-Objectives system was the difficulty in searching the database, and the need for technical help when doing any ad hoc search. Curriculum Index provides much simpler and more powerful search capabilities than the earlier system. It allows the user to specify searches of the database using any number of fields as search criteria. If multiple fields are specified, the system allows the user to indicate whether the fields should be combined using a Boolean AND (i.e. all selected fields must have matching values to constitute a matching record) or OR (i.e. any selected field with a matching value means the record matches). Similarly, in the case of fields which may have multiple data values, when more than one value is entered as a search

```
Record Number: 23      Form: Case Summary Sheet    Page: 1

Year of Course: 89-90

Course:          Human Body

Discipline:      Geriatrics
+                Gross Anatomy
                 Histology

Educational Format: Case Tutorial          Format Number: case2

Resources Used:  Books
+                Computer Programs

Author:          Johnson, Susan


Exercise Title:  The Case of the Falling Lawyer

Patient Sex: Female     Patient Age: 59     Patient Race: White

Press <TAB> to move forward or <SHIFT-TAB> to move backwards
```

Figure 1
Sample Data Entry Screen

Year of Course	(e.g. academic year 90-91)
Course(CV)	(name of course in which session is used)
Discipline(CV)	(academic field, e.g. oncology)
Educational Format(CV)	(e.g. Case Tutorial, Lab)
Format Number	(Identifying code number)
Resources Used(CV)	(e.g. slides, radiographs)
Author	(faculty who wrote this session)
Exercise Title	(Name of session)
Patient Sex(CV)	
Patient Age	
Patient Race(CV)	
Diagnoses	
Anatomical System(CV)	(e.g. cardiovascular system)
Anatomical Structure(CV)	(e.g. brain, heart)
Disease Process(CV)	(e.g. neoplasm, infection)
Pathophysiology(CV)	(e.g. arrhythmia, uremia)
Lab Tests(CV)	
Radiologic Studies(CV)	
Diagnostic Procedures(CV)	
Drugs(CV)	
Other Therapy(CV)	(e.g. dialysis, physical therapy)
Social Issues(CV)	(e.g. legal issues, public health issues)
Abstract	
References	
Other	

Note: items for which a controlled vocabulary has been specified are denoted by "(CV)" above.

Figure 2
Curriculum Index Data Items

criterion the user may indicate whether the values should be combined using AND or OR.

Once the set of records which match specified keyword entries has been identified, the user may browse through them on the screen or on a paper printout. Free-text fields such as **Abstract** and **References** can be examined to identify the records of interest for the current request. For instance, to find cases which deal with heart disease a user could specify that the **Anatomical System** field contain the entry "Cardiovascular System" and the **Anatomical Structure** field contain the entry "Heart". When the system presents all records which include these entries, the user can review the Abstract field to determine which sessions have heart disease as their primary subject and eliminate those which deal with normal, healthy hearts or in which heart disease is not the principal diagnosis.

In order to expedite the capture of data from faculty authors, the Curriculum Index system provides automated generation of data capture forms. The system prints a questionnaire showing each field to be entered for a record. For each field that has a controlled vocabulary, the vocabulary entries are shown in a checklist format. A sample page from the current capture form is shown in Figure 3. These forms, currently 6 pages long, are filled out by faculty who contribute new material for the curriculum. Since entries for many fields can be entered by simply checking off the appropriate items from the lists, these forms require relatively little time to complete. Also, since faculty are presented with the current vocabulary, the chance that they will select a predefined entry is strongly enhanced over a fill-in-the-blanks format. If a desired entry is not available and no existing entry will serve, the author notes this on the form. A faculty member reviews these terms periodically to determine which ones should be added to the controlled vocabulary. Since the generation of the forms is done by the Curriculum Index system itself, the most current vocabulary is always provided when a new form is printed.

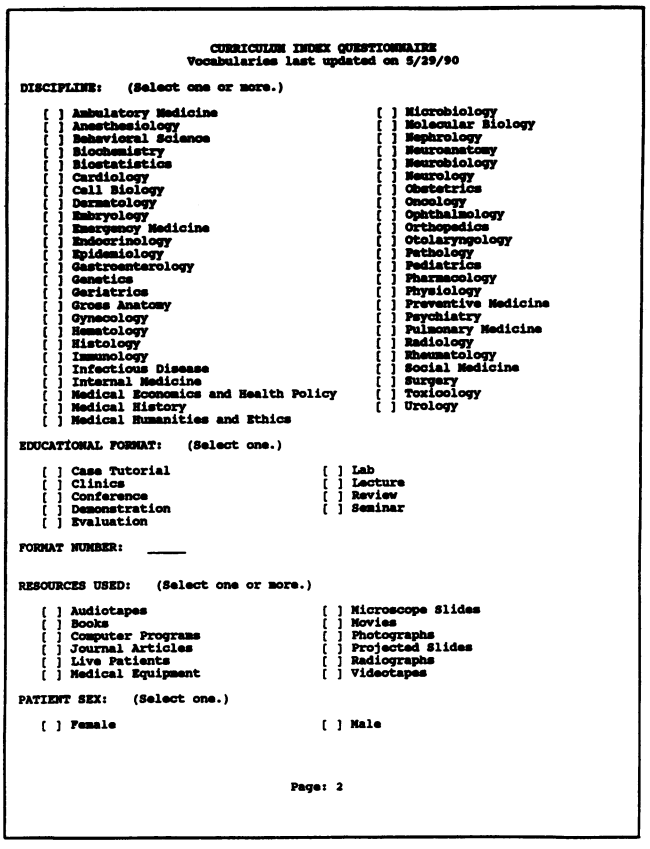
Current Status

As of August, 1990, all of the 55 cases used in the first-year courses have been entered into the system. To index these cases, about 20 terms needed to be added to the original controlled vocabularies. No additional fields were needed. The data entry person has been able to do all desired searches of the database without consultation with technical staff. Faculty authors have been generally positive on the use of the checklist as a way of capturing data to be entered into the system.

The initial system provided only very basic reporting capabilities, under the assumption that the best way to find out what types of reports would be needed was to try the system out in practice. In the first six months of system operation, several additional options for printout of retrieved information have been implemented. Now that a substantial database of information has been entered into the system, publicity is underway to make faculty aware that the system is available to them as a resource. The system has already been used to respond to several faculty inquiries about where different topics are addressed in the curriculum.

Future Plans

We expect to continue to enter data into the Curriculum Index system as new course material is produced by faculty and as teaching sessions evolve. Current plans include entry of data on lectures and laboratory sessions for the first-year courses as well as the cases used in the second year courses. Since about 25% of the cases are replaced or reworked each year, additional case information for first-year courses will also be captured as new cases are generated. We plan to add to the controlled vocabular-



CURRICULUM INDEX QUESTIONNAIRE
Vocabularies last updated on 5/29/90

DISCIPLINE: (Select one or more.)

[ ] Ambulatory Medicine
[ ] Anesthesiology
[ ] Behavioral Science
[ ] Biochemistry
[ ] Biostatistics
[ ] Cardiology
[ ] Cell Biology
[ ] Dermatology
[ ] Embryology
[ ] Emergency Medicine
[ ] Endocrinology
[ ] Epidemiology
[ ] Gastroenterology
[ ] Genetics
[ ] Geriatrics
[ ] Gross Anatomy
[ ] Gynecology
[ ] Hematology
[ ] Histology
[ ] Immunology
[ ] Infectious Disease
[ ] Internal Medicine
[ ] Medical Economics and Health Policy
[ ] Medical History
[ ] Medical Humanities and Ethics
[ ] Microbiology
[ ] Molecular Biology
[ ] Nephrology
[ ] Neuroanatomy
[ ] Neurobiology
[ ] Neurology
[ ] Obstetrics
[ ] Oncology
[ ] Ophthalmology
[ ] Orthopedics
[ ] Otolaryngology
[ ] Pathology
[ ] Pediatrics
[ ] Pharmacology
[ ] Physiology
[ ] Preventive Medicine
[ ] Psychiatry
[ ] Pulmonary Medicine
[ ] Radiology
[ ] Rheumatology
[ ] Social Medicine
[ ] Surgery
[ ] Toxicology
[ ] Urology

EDUCATIONAL FORMAT: (Select one.)

[ ] Case Tutorial
[ ] Clinics
[ ] Conference
[ ] Demonstration
[ ] Evaluation
[ ] Lab
[ ] Lecture
[ ] Review
[ ] Seminar

FORMAT NUMBER: ____

RESOURCES USED: (Select one or more.)

[ ] Audiotapes
[ ] Books
[ ] Computer Programs
[ ] Journal Articles
[ ] Live Patients
[ ] Medical Equipment
[ ] Microscope Slides
[ ] Movies
[ ] Photographs
[ ] Projected Slides
[ ] Radiographs
[ ] Videotapes

PATIENT SEX: (Select one.)

[ ] Female
[ ] Male

Page: 2

Figure 3
Sample Page from Data Capture Form

ies as required to index these additional sessions, as well as providing any additional retrieval and reporting capabilities that prove to be needed. One of the advantages of using a general-purpose database system such as INFOMAN as a base for this system is that changes like these are easy to make.

At present all retrieval questions are handled by the data entry person. However, the system was designed to be easy enough to learn and use that it could be made available to interested faculty and curriculum coordinators for their reference. In the future, we also hope to make it available to students as one component of a student personal reference file. The database would be useful to students who want to find out where in their studies they encountered a certain diagnosis or anatomical system; as noted above, in a case-based format such as the one now in use at Harvard, these questions are not always readily answered without a resource such as the Curriculum Index.

Conclusion

Harvard Medical School has had almost five years of experience with the development and use of computer systems for monitoring the content of the medical curriculum. The ongoing changes in the HMS curriculum and the use of the case-based format make a curriculum indexing system particularly important to insure that all needed content is covered. Our experience has led us toward the use of a generalized database management system that offers easy modification of data entry forms and controlled vocabularies as well as powerful searching capabilities. The importance of using controlled vocabularies has been borne out, but the MeSH vocabulary was found to be inappropriate for this application in our environment. We plan to continue our efforts in this area in the future, and hope to make our database available not only to curriculum coordinators but to faculty and students.

Acknowledgements

The work described in this paper was supported in part by grants from Hewlett-Packard Corporation and from the Henry J. Kaiser Family Foundation. MUMPS is a trademark of the Massachusetts General Hospital. DTM-PC is a trademark of DataTree, Inc.

References

[1] Tosteson DC. New pathways in general medical education. NEJM 1990; 322:234-238.

[2] Feletti GI, Carver ST (eds.). The new pathway to general medical education at Harvard University. Teaching and Learning in Medicine 1989; 1:42-46.

[3] Barnett GO. Information technology and undergraduate medical education. Academic Medicine April, 1989; pp. 187-190.

[4] Barnett GO, Piggins JL, Moore GT, Foster EA. Information technology and medical education. Hewlett-Packard Journal 1987 38:28-35.

[5] Timson G. VA FileMan User's Manual. Publication of the MUMPS Users' Group, 4321 Hartwick Rd, College Park Md 20740.

[6] Bowie J, Barnett GO. MUMPS - an economical and efficient time-sharing system for information management. Comp Prog Biomed 1976 6:11-22.

[7] Medical subject headings. In Cumulated Index Medicus [NIH Publication No. 84-259] 1983 24:445-446.

[8] Gotlib D, DesGroseillers JP, Dolphin, P. A computerized database-management system for curriculum analysis. Can Med Assoc J 1984 131:861-863.

[9] Buckenham S, Sellers EM, Rothman AI. An application of computers to curriculum review and planning. J Med Educ 1986 61:41-45.

[10] Dimse SS, O'Connell MT. Cataloging a medical curriculum using MeSH keywords. In: Greenes (ed.) Proc Twelfth Annual Symposium on Computer Applications in Medical Care (IEEE Press, New York, 1988) 332-336.